

doi: 10.3969/j.issn.1674-1242.2022.04.004

# 精神分裂症组蛋白修饰与神经递质系统 关系的多组学研究

闻祎韡, 宋炜宸

(上海交通大学生物医学工程学院, 上海 200030)

**【摘要】** 目的 探索组蛋白修饰通过神经递质系统影响精神分裂症发病风险的途径。方法 利用 WGCNA 软件包对精神分裂症患者脑表达谱建立的组蛋白 - 神经递质间的基因共表达网络, 分析患者和健康对照组是否在组蛋白 - 神经递质的关联关系上存在差异。结果 患者大脑中组蛋白 - 神经递质系统相关性显著降低。结论 通过辨别网络中疾病差异表达和新发突变的核心组蛋白修饰基因的分布, 该研究发现 CALM2、PRKAR1A、CAPZA2、BSN、GRIN1、LRFN3 六个候选基因可能是精神分裂症遗传风险的关键基因。

**【关键词】** 精神分裂症; 组蛋白; 神经递质系统; CALM2; 风险基因

**【中图分类号】** R749

**【文献标志码】** A

文章编号: 1674-1242(2022)04-0219-04

## Multi-omics Study of the Relationship Between Histone Modification and Neurotransmitter System in Schizophrenia

WEN Yiwei, SONG Weichen

(Shanghai Jiao Tong University, School of BME, Shanghai 200030, China)

**【Abstract】 Objective** To explore the pathway through which histone modification affects the risk of schizophrenia through the neurotransmitter system. **Methods** To use the WGCNA package to establish the gene co-expression network between histone and neurotransmitter in the brain expression profiles of patients with schizophrenia, to analyze whether there are differences in histone-neurotransmitter associations between patients and healthy controls. **Results** The correlation of histone-neurotransmitter system in the patient's brain was significantly reduced. **Conclusion** By identifying the distribution of core histone-modifying genes with differential expression and newly mutated core histones in the network, we found CALM2, PRKAR1A, CAPZA2, BSN, GRIN1, LRFN3 Six candidate genes may be key genes for genetic risk of schizophrenia.

**【Key words】** Schizophrenia; Histone; Neurotransmitter; CALM2; Risk Gene

### 1 背景

精神分裂症 (Schizophrenia, SCZ) 是比较常见和严重的精神疾病之一, 使患者与患者家属的日常生活受到严重影响。在全球疾病总负担中, 约 8% 来自精神疾病, 在全球非致命疾病中, 精神分裂症位列第七<sup>[1]</sup>。关于精神分裂症的病因,

目前还没有定论, 但学者们一般都认同精神分裂症的发生是遗传和环境两个因素共同作用的结果。过往很多研究都证实了遗传因素对精神分裂症有重要影响, 精神分裂症的遗传度很高<sup>[2]</sup>。其中, 神经递质紊乱假说认为精神分裂症的发生与神经递质的功能异常有很大的相关性。具体而言, 神经递质紊乱假说包括多巴胺 (Dopamine, DA) 功

收稿日期: 2022-11-16

作者简介: 闻祎韡 (2000—), 女, 吉林省吉林市人, 本科生, 从事精神疾病遗传学研究。E-mail: wenyiwei@sjtu.edu.cn

通信作者: 宋炜宸, 男, 博士, 电话 (Tel.): 18317185227; E-mail: goubegou@sjtu.edu.cn

能亢进学说、5-羟色胺(5-HT)功能紊乱学说、谷氨酸功能紊乱学说等,其中存在时间最久也最受重视的是多巴胺功能亢进学说<sup>[3]</sup>。在全基因组关联研究(Genome-Wide Association Studies, GWAS)中得到的108个精神分裂症相关位点中,大部分位点都与多巴胺、谷氨酸等神经递质相关。

组蛋白修饰是表观遗传的重要组成部分,表观遗传是指在DNA序列不变的情况下,出现可逆、可遗传的遗传物质的变化,包括DNA甲基化、组蛋白修饰、非编码RNA等。其中,针对DNA甲基化与精神疾病的相关性的研究最多,并且有充分的证据证明其与精神疾病的发生相关<sup>[4]</sup>。目前,关于组蛋白修饰与精神分裂症和孤独症的研究还较少,缺少有力的证据能直接证明两者的相关性,相关机制也有待进一步研究。但最早的证据是在临床上发现丙戊酸钠——一种组蛋白去乙酰化酶(Histone Deacetylase, HAT)抑制剂,能够缓解精神分裂症的症状<sup>[5]</sup>。我们推测,组蛋白修饰对神经递质的调控异常与精神疾病的发生可能存在某种关系。为了阐明组蛋白修饰可以通过影响神经递质系统来影响疾病的发生,本研究通过基因组学和转录组学融合的方式,把精神分裂症和孤独症谱系障碍作为研究对象,通过研究组蛋白-神经递质关联性与精神疾病的关系,研究组蛋白修饰与精神疾病的关系,并进一步基于该假设,寻找核心的组蛋白修饰基因,建立疾病的组蛋白-神经递质核心调控网络,从中寻找新的候选基因。

## 2 方法

### 2.1 数据收集与预处理

本研究用到的具体数据有谷氨酸能突触基因数据、染色质共价修饰基因数据、多巴胺基因数据、谷氨酸受体复杂基因数据<sup>[6]</sup>、5-羟色胺受体基因数据<sup>[7]</sup>、614名健康志愿者与286名精神分裂症患者的海马体<sup>[8]</sup>和背外侧前额叶皮质基因数据,以及转录组测序数据(RNA-seq)的FPKM(Fragments Per Kilobase Per Million Mapped Reads)数据。

为了减少数据运算量、提高计算效率,需要筛选掉在各个样本中表达差异不大的基因,只保留方差在前25%的基因。利用WGCNA软件包中的goodSamplesGenes()函数,该函数能迭代地识别有过多缺失条目的样本及零方差的基因,用以确认数据质量,并可据此剔除不合格的样本和基因数据。对样本

聚类,绘制聚类树,根据其形貌剔除离群的样本。

### 2.2 加权基因共表达网络分析<sup>[9]</sup>

加权基因共表达网络分析(Weighted Gene Co-expression Network Analysis, WGCNA)的核心步骤:首先,计算任意两个基因之间的相关系数,获得相似性矩阵;其次,根据选择的软阈值对相关系数进行加权从而获得邻接矩阵;再次,利用特定算法获得TOM矩阵,TOM矩阵能纳入非直接的基因相关性信息,可以更加精确地描述基因表达谱的类似性;最后,对TOM矩阵进行动态树分割,划分出基因模块,并合并出高度相似的模块,一般认为同一模块中的基因具有相似的表达谱。通过pickSoftThreshold()函数得到软阈值后,可以直接采用blockwiseModules()函数一步法构建共表达网络。在参数设置上,将计算机一次性处理的最大模块的基因数设置为5000,最小模块的基因数设置为30,合并模块的阈值设置为0.25。这个函数能合并出较高相似性的模块,具体则要通过计算每个模块的模块特征基因来合并相似的模块,并用plotDendroAndColors()函数绘制图像,观察所构建的共表达网络的聚类树与各个模块的颜色情况。再利用TOMsimilarityFromExpr()函数单独获得TOM矩阵,供后续分析使用。在患者组与健康对照组之间,通过t检验比较TOM矩阵参数的高低,以分析患者大脑中组蛋白与神经递质系统之间的关联是否出现变化。

### 2.3 建立疾病的组蛋白-神经递质调控核心网络

从PsymuKB数据库中寻找精神分裂症患者中携带严重新发变异(De Novo Variants, DNVs)的染色质修饰基因<sup>[10]</sup>。在精神分裂症患者的组蛋白-神经递质共表达网络中,从与组蛋白修饰基因具有高相关性的神经递质基因中,寻找这些新发突变的基因。我们定义与组蛋白高度相关的基因为:该神经递质基因与组蛋白修饰基因的TOM阈值为所有神经递质基因与该组蛋白修饰基因的前1%。综合上述结果得到精神分裂症的关键基因,回到精神分裂症的组蛋白-神经递质共表达网络中,选择合适的TOM阈值,提取与这些关键基因具有高相关性的基因,与关键基因共同组成大小适中的核心网络。将这些基因及其相互之间的TOM阈值信息整理成Cytoscape<sup>[11]</sup>可以识别的格式,绘制网络图。利用Cytoscape的Network Analyzer工具,找到连通性最高的几个基因,突出显示,作为候选基因。

### 3 结果

通过 WGCNA 软件包分析, 患者组的基因最终分成了 27 个基因模块, 健康对照组的基因最终分成了 22 个基因模块。我们统计了组蛋白与各个神经递质系统基因的 TOM 阈值, 发现患者组 TOM 阈值均显著低于健康对照组, 如表 1 所示。

表 1 精神分裂症患者组蛋白与神经递质系统之间的关联异常

Tab. 1 Difference of histone modification-neurotransmitter correlation between schizophrenia patients and healthy control

多巴胺系统	患者组	健康对照组
mean	$3.434 \times 10^{-3}$	$6.579 \times 10^{-2}$
p-value		$2.20 \times 10^{-16}$
t		-36.6
df		3 830
5-HT 系统	患者组	健康对照组
mean	$2.661 \times 10^{-2}$	$8.600 \times 10^{-2}$
p-value		$2.20 \times 10^{-16}$
t		-48.5
df		13 000
谷氨酸系统	患者组	健康对照组
mean	$1.555 \times 10^{-2}$	$7.332 \times 10^{-2}$
p-value		$2.20 \times 10^{-16}$
t		-237
df		245 000
综合结果	患者组	健康对照组
mean	$1.615 \times 10^{-2}$	$7.393 \times 10^{-2}$
p-value		$2.20 \times 10^{-16}$
t		-251
df		282 000

在构建精神分裂症核心网络时, 所选用的 TOM 阈值为 0.15, 得到了一个包括 225 个节点、909 条边的网络。如图 1 所示, 在所得到的精神分裂症核心网络中, CALM2、PRKAR1A、CAPZA2、BSN、GRIN1、LRFN3 这六个基因的连通性最高, 可以作为候选基因。

图中, CALM2 基因 (Calmodulin 2) 是编码钙调蛋白的基因之一, 钙调蛋白是一种参与细胞内信号传导的钙结合蛋白, 可通过对钙离子的调节, 参与调节细胞间的信号传导, 细胞的运动、分化和增殖, 神经递质的合成、运输和释放等活动。CALM2 基因与常染色体显性儿茶酚胺能多态性室性心动过速 (Catecholaminergic Polymorphic Ventricular Tachycardia, CPVT) 和长 QT 综合征 (Long QT Syndrome, LQTS) 相关<sup>[12]</sup>。

### 4 讨论

回顾以往的研究结果<sup>[13]</sup>, 有多重证据证明了神经递质与精神疾病的相关性, 以及基因表观遗传

与精神疾病的相关性, 但目前基因表观遗传研究中对组蛋白修饰的研究还比较少。在本研究中, 我们基于 WGCNA 软件包建立了精神分裂症、健康对照组的组蛋白-神经递质共表达网络, 通过分析各组的组蛋白-神经递质 TOM 矩阵、组蛋白-神经递质共表达基因对组蛋白-神经递质共表达基因的连接程度, 多层次、多角度比较了精神疾病和健康对照组的组蛋白-神经递质关联性差异。结果表明, 组蛋白-神经递质关联性在健康对照组与患者组之间存在差异, 并且总体上在患者组中关联性降低。在精神分裂症中, 分析结果都呈阳性, 即在精神分裂症中组蛋白-神经递质关联性下降, 提示组蛋白修饰对神经递质的调控出现异常, 影响了疾病的发生<sup>[14]</sup>, 具体机制还有待更多的样本数据与进一步研究加以阐释。

而后基于先前所建立的精神疾病的组蛋白-神经递质共表达网络及公开数据库中的精神分裂症中的差异表达基因和新发突变基因数据, 找到精神分裂症中差异表达的核心组蛋白修饰基因 30 个, 精神分裂症中新发突变的核心组蛋白修饰基因 25 个。根据这些基因, 回到精神疾病的组蛋白-神经递质 TOM 矩阵, 选择合适的阈值, 找到与这些基因具有高相关性的基因, 共同组成精神疾病的组蛋白-神经递质核心网络, 通过 Cytoscape 软件绘制核心网络图, 并利用其中的分析工具找到连通性最高的几个基因作为候选基因, 最后在精神分裂症的核心网络中, 找到了 CALM2、PRKAR1A、CAPZA2、BSN、GRIN1、LRFN3 这六个候选基因, 可以为未来针对精神分裂症和孤独症的发病机制与诊断靶点的研究提供新的思路。

本研究的结果目前仅是初步结论, 可以提供对未来疾病研究的建议, 若要从组蛋白修饰对神经递质的调控角度阐明精神疾病发病的机制, 还有待进一步的实验研究与分析, 也期待进一步的研究推动发病机制的发展, 早日找到有效治疗靶点, 减轻精神疾病所带来的家庭和社会负担。此外, 精神分裂症等精神疾病受到遗传因素和环境因素的共同影响, 存在区域差异性, 而本研究用于分析的样本数据量有限, 要获得真实有效的基因靶点还需要对不同地区、民族的人的数据进行分析, 本研究的结果存在一定局限性。

#### 参考文献

- [1] MURRAY C J, LOPEZ A D. Alternative projections of mortality

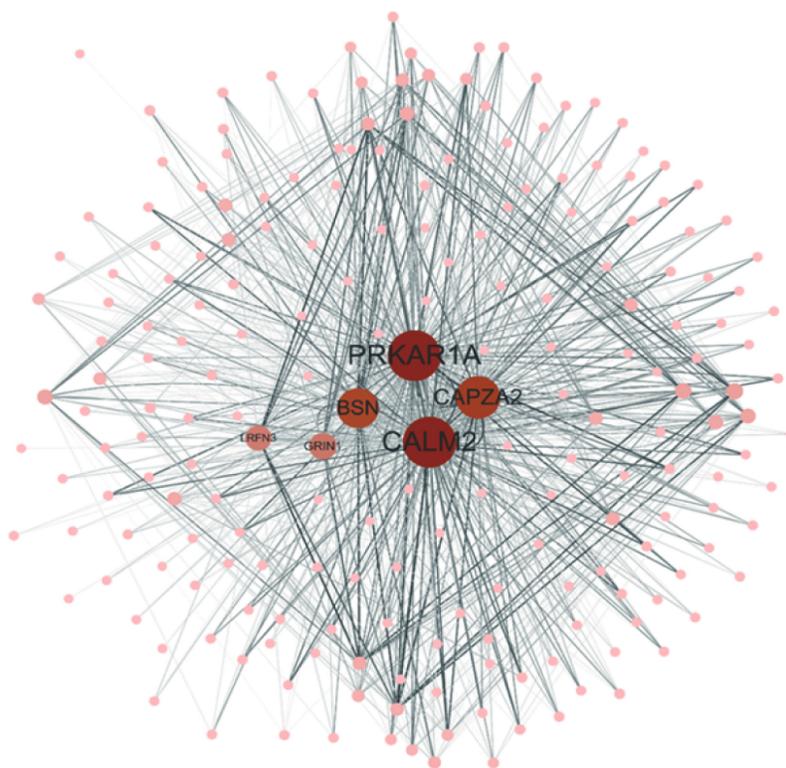


图1 精神分裂症的核心网络

Fig.1 Core regulation network of schizophrenia

- and disability by cause 1990–2020: Global Burden of Disease Study [ J ] . **Lancet**, 1997.349(9064): 1498–504.
- [ 2 ] SULLIVAN P F, AGRAWAL A, BULIK C M, *et al.* Psychiatric Genomics: An Update and an Agenda [ J ] . **Am. J. Psychiatry**, 2018.175(1): 15–27.
- [ 3 ] HOWES O D, MCCUTCHEON R, OWEN M J, *et al.* The Role of Genes, Stress, and Dopamine in the Development of Schizophrenia [ J ] . **Biol. Psychiatry**, 2017.81(1): 9–20.
- [ 4 ] HANNON E, DEMPSTER E, VIANA J, *et al.* An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation [ J ] . **Genome Biol.**, 2016.17(1): 176.
- [ 5 ] SHARMA R P, GRAYSON D R, GAVIN D P. Histone deacetylase 1 expression is increased in the prefrontal cortex of schizophrenia subjects: analysis of the National Brain Databank microarray collection [ J ] . **Schizophr. Res.**, 2008.98(1–3): 111–117.
- [ 6 ] KIROV G, POCKLINGTON A J, HOLMANS P, *et al.* De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia [ J ] . **Mol. Psychiatry**, 2012.17(2): 142–53.
- [ 7 ] PARDINAS A F, HOLMANS P, POCKLINGTON A J, *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection [ J ] . **Nat. Genet.**, 2018.50(3): 381–389.
- [ 8 ] JAFFE A E, STRAUB R E, SHIN J H, *et al.* Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis [ J ] . **Nat. Neurosci.**, 2018.21(8): 1117–1125.
- [ 9 ] LANGFELDER P, HORVATH S. WGCNA: an R package for weighted correlation network analysis [ J ] . **BMC Bioinformatics**, 2008.9: 559.
- [ 10 ] LIN G N, GUO S, TAN X, *et al.* PsyMuKB: An Integrative De Novo Variant Knowledge Base for Developmental Disorders [ J ] . **Genomics. Proteomics Bioinformatics**, 2019.17: 453–464.
- [ 11 ] SHANNON P, MARKIEL A, OZIER O, *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks [ J ] . **Genome Res.**, 2003.13: 2498–2504.
- [ 12 ] CROTTI L, JOHNSON C N, GRAF E, *et al.* Calmodulin mutations associated with recurrent cardiac arrest in infants [ J ] . **Circulation**, 2013.127(9): 1009–1017.
- [ 13 ] The Network and Pathway Analysis Subgroup of the Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways [ J ] . **Nat. Neurosci.**, 2015.18(2): 199–209.
- [ 14 ] FAUNDES V, NEWMAN W G, BERNARDINI L, *et al.* Histone Lysine Methylases and Demethylases in the Landscape of Human Developmental Disorders [ J ] . **Am. J. Hum. Genet.**, 2018.102(1): 175–187.